Promotion Time Cure Rate Model with Random Effects: an Application to a Multi-centre Clinical Trial of Carcinoma

Diego I. Gallardo*[1], Heleno Bolfarine[2], Antonio C. Pedroso-de-Lima[3]

Instituto de Matemática e Estatística

Universidade de Sao Paulo, Brazil

*diegog@ime.usp.br; [2]hbolfar@gmail.com; [3]acarlos@ime.usp.br

Abstract

This paper extends the cure rate model considered in Lopes and Bolfarine (2012) by taking into account random effects in both, the probability of cure and the survival function for individuals that are at risk. The model is parametrized in terms of the cured fraction which is then linked to covariates. The estimation is based on the restricted maximum likelihood (REML) approach proposed in McGilchrist and Yau (1995). Simulation studies are performed and results based on a real data set are presented indicating good performance of the proposed approach.

Keywords

Promotion Time Cure Rate Model; Random Effects; REML

Introduction

An implicit assumption with the ordinary survival models is that all individuals under study are susceptible to the event of interest, which is not always in the case given the improvement in disease treatments experienced in the last decades. For some types of cancer, for example, great improvement has increased the probability that an individual is considered with the disease under control (usually denoted by a cure). The amount of cured individuals after a treatment is typically known as the cure fraction. An indication of the presence of cured individuals in a specific treatment is given by the value that the Kaplan-Meier estimator takes at the maximum survival time. A formal test for the presence of a cure fraction using the Kaplan-Meier estimator can be seen in Maller and Zhou (1996).

Berkson and Gage (1952) developed a model that came to be known in the literature as the mixture model for cure fraction where there is a proportion p of susceptible individuals and, hence, a proportion $1-p$ of cured individuals. This model has been studied by several authors including Goldman (1984), Sposto et al. (1992) and Mizoi et al. (2007), among others. An alternative route was pursued by Yakolev and Tsodikov (1996) and Chen et al. (1999). Their approach is based on the assumption that each individual has a fixed unobserved (latent) number N of factors, each capable of triggering the event of interest. This model, presented in Section 2, known in the literature as the promotion time cure rate model has been the subject of intense research activity. Further references can be found in Lopes and Bolfarine (2012).

Some literature connecting and unifying both, the mixture and promotion time models are Yin and Ibrahim (2005) and Rodrigues et al. (2009)

This paper aims to consider situations where the collected data come from several clusters and are obtained using repeated measures. For example, in multicenter clinical trials, patients from a given clinic (cluster) can be affected by unobserved variables specific to that particular center. Therefore, it seems appropriate to include in the model random effects to take into account the dependence generated in patients treated in a same clinic. Yau and Ng (2001), working with a mixture model incorporates random effects in both, the cure fraction (V_j) and on the survival function of the susceptible individuals (U_j), assuming that the random effects are normally and independently distributed. Lai and Yau (2008) worked with the bivariate normal to (U_j, V_j) in the mixture model, assuming a possible dependence between the two random effects in a same clinic. However, either was developed a classical approach to parameter estimation. Lopes and Bolfarine (2012) incorporated random effects in the promotion time cure rate model only on the cure fraction and developed classical as well as Bayesian approaches for parameter estimation.

In this paper, the model in Lopes and Bolfarine (2012) is extended to develop a classical approach to the parameter estimation in the promotion time cure rate model incorporating random effects for both, the survival time of susceptible individuals and in the cure fraction.

Section 2 presents the ordinary promotion time cure rate model. In Section 3, the model is extended, by incorporating random effects. In Section 4, the classical approach to parameter estimation is developed, based on the BLUP and REML estimation approaches. In Section 5, an application of the proposed model is considered for a real data set. In Section 6, a simulation study is presented to evaluate the proposed estimation procedures. Final considerations and discussions are presented in Section 7, where it is concluded that the proposed estimation approach is fairly satisfactory and that the model is a viable alternative to fit real data where it seems adequate to utilize random effects along with a cure fraction.

The Promotion Time Cure Rate Model (PTCRM)

The model proposed in Chen et al. (1999) can be defined in the following way. Suppose that an individual in the population has N carcinogenic cells at the time entering the study, where $N \sim Po(\theta)$, $\theta > 0$. Moreover, W_a is denoted as the random variable expressing the time the a-th cell needs to produce a detectable cancer. For non-cured subjects, there are cells such as $N > 0$, with W_a, $a = 1,2,\dots,N$, conditionally independent given N and identically distributed with common survival function $F(t|\boldsymbol{\kappa}) = 1 - S(t|\boldsymbol{\kappa})$, where $\boldsymbol{\kappa}$ is a set of unknown parameters, properly defined in some parameter space. For cured subjects, $N = 0$ and it is assumed that $P(W_0 = \infty) = 1$. The distribution F is related to the susceptible individuals and, in general, it is a proper function in the sense that $\lim_{t\to\infty} F(t|\boldsymbol{\kappa}) = 1$. Under this framework, the time for a cancer to be detected can be represented by a random variable $T = \min\{W_a,\ 0 \le a \le N\}$. In a word, for non-cured individuals, the failure time is the minimum among the times the cells may take to, eventually, produce a detectable cancer whereas cured individuals will never experience the event of interest and the failure time in this case is infinity. Under such conditions, it can be verified that the (improper) survival function for the random variable T, also called the population survival function, is given by

$$S_p(t) = e^{-\theta F(t|\boldsymbol{\kappa})}. \quad (1)$$

Therefore, the cure fraction in the population is given by $\lim_{t\to\infty} S_p(t) = e^{-\theta} > 0$. Assuming right censoring, the actual data for the i-th subject may be represented by the random variables $Y_i = \min(T_i, C_i)$ and $\delta_i = I(T_i \le C_i)$, where T_i and C_i are failure and censoring times, respectively, for $i = 1,\dots,n$. The observed data will be denoted by $\boldsymbol{D}_{obs} = (\boldsymbol{y}, \boldsymbol{\delta})$, where $\boldsymbol{y} = (y_1,\dots,y_n)$ and $\boldsymbol{\delta} = (\delta_i,\dots,\delta_n)$. It is noted that $\boldsymbol{N} = (N_1,\dots,N_n)$ are non-observable, and the complete data will be denoted by $\boldsymbol{D} = (\boldsymbol{y}, \boldsymbol{\delta}, \boldsymbol{N})$. Therefore, the complete likelihood function for $(\boldsymbol{\kappa}, \theta)$ can be written as

$$L(\boldsymbol{\kappa}, \theta|\boldsymbol{D}) = \left(\prod_{i=1}^{n} S(y_i|\,\boldsymbol{\kappa})^{N_i-\delta_i}[N_i f(y_i|\boldsymbol{\kappa})]^{\delta_i}\right) \times \exp\{\sum_{i=1}^{n}[N_i \log(\theta) - \log(N_i!)] - n\theta\}. \quad (2)$$

The marginal likelihood with respect to $\boldsymbol{N}$ is [Chen et al. (1999)]

$$L(\boldsymbol{\kappa}, \theta|\boldsymbol{D}_{obs}) = \prod_{i=1}^{n}\big(\theta f(y_i|\boldsymbol{\kappa})\big)^{\delta_i}\exp\big\{-\theta\big(1 - S(y_i|\boldsymbol{\kappa})\big)\big\}. \quad (3)$$

Consider now the particular case where the lifetimes W_a, $a = 1, 2, ...$, have a Weibull distribution. Although other distributions could also be considered, the Weibull has a number of advantages: it is popular, easy to interpret and has the flexibility of accommodating different shapes for the hazard function, depending on the shape parameter α. Moreover, for the practical situation motivating the model, it is expected $\alpha > 1$.

For non-homogeneous populations, let $\boldsymbol{x}_i = (x_{i1},\dots,x_{ir})^T$ be a covariate vector associated with the lifetimes $W_a, a = 1,2,....$ Following Yau and Ng (2001) is considered

$$S(t|\xi_i, \boldsymbol{\kappa}) = [S_0(t|\boldsymbol{\kappa})]^{\exp\,\xi_i},\ \ i = 1,\dots,n, \quad (4)$$

where $S_0(t|\boldsymbol{\kappa}) = \exp\{-\lambda t^{\alpha}\}$, $\boldsymbol{\kappa} = (\lambda, \alpha)$ and $\xi_i = \boldsymbol{x}_i^T\boldsymbol{\gamma}$.

As for the cure fraction, a vector of covariates $\boldsymbol{z}_i = (z_{i1},\dots,z_{is})^T$ may be assumed. Consider

$$\theta_i = \exp(\boldsymbol{z}_i^T\boldsymbol{\beta}),\ \ i = 1,\dots,n.$$

The quantities $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$ are, respectively, $r \times 1$ and $s \times 1$ vectors of unknown parameters. The choice of covariates for each part depends on the practical problem being considered and on the information given by the physician.

Note that the linear function $\boldsymbol{x}_i^T\boldsymbol{\gamma}$ should exclude an intercept because if that is the case then,

$$S(t|\xi_i, \boldsymbol{\kappa}) = [S_0(t|\boldsymbol{\kappa})]^{\exp\,\xi_i} = [S_0(t|\boldsymbol{\kappa})]^{\exp\{\gamma_0 + \sum_{k=1}^{r} x_{ik}\gamma_k\}} = [S_0(t|\boldsymbol{\kappa}^*)]^{\exp\{\sum_{k=1}^{r} x_{ik}\gamma_k\}}, \quad (5)$$

which will lead to identifiability problems, as

mentioned in Yau and Ng (2001).

Promotion Time Cure Rate Model with Random Effects

Suppose n_j patients grouped in clinic j, $j = 1, \ldots, J$. The expanded notation is considered in the previous section by means of the proposition that N_{ij} represents the number of carcinogenic cells associated with the i-th individual from the j-th clinic, with $N_{ij} \sim Po(\theta_{ij})$. For $j = 1, \ldots, J$, define the vector $\boldsymbol{N}_j = (N_{1j}, N_{2j}, \ldots, N_{n_j j})^T$. Let W_{aij} be the random variable corresponding to the time the a-th carcinogenic cell from the i-th individual within the j-th clinic needs to yield a detectable cancer. For a given subject, the random variables $W_{aij}, a = 1, \ldots, N_{ij}$, conditionally independent given N_{ij} have a common survival $S(t|\xi_{ij}, \boldsymbol{\kappa})$. Hence, the time up to a detectable cancer can be defined as $T_{ij} = \min\{W_{aij}, 1 \leq a \leq N_{ij}\}$ for non-cured individuals and $P(T_{ij} = \infty) = 1$ if $N_{ij} = 0$. The survival function $S(\cdot\,|\xi_{ij}, \boldsymbol{\kappa})$ is the same as in (4), with covariates $\boldsymbol{x}_{ij}$ and $\boldsymbol{z}_{ij}$ included in the model in the same way as discussed in the previous section. However, in order to take into account the effect of the j-th clinic on survival for cured and uncured subjects, random effects $(U_j, V_j), j = 1, \ldots, J$ are considered so that

$$\xi_{ij} = \boldsymbol{x}_{ij}^T \boldsymbol{\gamma} + U_j \quad \text{and} \quad \theta_{ij} = \exp\{\boldsymbol{z}_{ij}^T \boldsymbol{\beta} + V_j\}, \tag{6}$$

where $(U_j, V_j) \sim N_2(\boldsymbol{0_2}, \boldsymbol{\Sigma}), j = 1, \ldots J$, with $\boldsymbol{\Sigma} = \text{diag}(\sigma_u^2, \sigma_v^2)$. For a given clinic j, high (low) values of U_j will lead to an increase (decrease) in the risk to activate a particular cancer cell; as a consequence subjects from that clinic will experience a greater (smaller) risk of developing a carcinogenic tumor. Similarly, low (high) values of V_j indicate that patients treated in the j-th clinic have greater (smaller) chance of being cured.

The model in Lopes and Bolfarine (2012) is a particular case of (6), when $V_j \sim N(0, \sigma_v^2), j = 1, \ldots, J$ and $\sigma_u^2 = 0$.

It should be mentioned here that the same non-identifiability issues discussed in Yau and Ng (2001) may arise, specially in trials with very long follow-up times, when it is difficult to distinguish between cure and censoring for long-term survivors.

Given that a non-intercept model is considered, the random effects $U_j, j = 1, \ldots, J$ will represent those quantities in this new formulation and need to be estimated. The solution proposed in Yau and Ng (2001) may be considered: the remaining parameters are estimated assuming the quantity $S_0(t|\,\boldsymbol{\kappa})$ is known. Define $\boldsymbol{U} = (U_1, \ldots, U_J)^T$ and $\boldsymbol{V} = (V_1, \ldots, V_J)^T$. The complete log-likelihood function for the model (after summing over $\boldsymbol{N} = (N_1, \ldots, N_J)$) is given by $l = l_1 + l_2$, in which

$$l_1 = \sum_{j=1}^{J} \sum_{i=1}^{n_j} \{\delta_{ij} [\log \theta_{ij} + \log f(t_{ij}|\boldsymbol{\kappa})] - \theta_{ij} F(t_{ij}|\boldsymbol{\kappa})\},$$

$$l_2 = (-1/2)\left[J \log(2\pi\sigma_u^2) + \left(\frac{1}{\sigma_u^2}\right) \boldsymbol{U}^T \boldsymbol{U}\right]$$

$$+(-1/2)\left[J \log(2\pi\sigma_v^2) + \left(\frac{1}{\sigma_v^2}\right) \boldsymbol{V}^T \boldsymbol{V}\right],$$

with $F(t|\xi_i, \boldsymbol{\kappa}) = 1 - [S_0(t|\boldsymbol{\kappa})]^{\exp\,\xi_i}$ and $S_0(t|\boldsymbol{\kappa}) = \exp\{-\lambda t^{\alpha}\}$.

Estimation

In the previous section it was observed that the log-likelihood function l may be expressed as a sum of two quantities l_1 and l_2, where due to the unobserved quantities $\boldsymbol{U}$ and $\boldsymbol{V}$, it can not be directly maximized. One possible way to deal with this problem is to consider the quadrature procedure discussed in Davis and Rabinowitz (1975). Typically, this approach requires high computation times since in the situation discussed here there are $2J$ random effects to deal with; moreover, there is the question of how many knots to consider, which can significantly affect the estimation procedure.

Another alternative is to consider the restricted maximum likelihood (REML) approach proposed by McGilchrist and Yau (1995) as well as Yau and Ng (2001), Lai and Yau (2008) and Lopes and Bolfarine (2012). Initially, define $\boldsymbol{\zeta} = (\boldsymbol{\gamma}, \boldsymbol{\beta}, \boldsymbol{U}, \boldsymbol{V})$. Assuming that $S_0(t|\boldsymbol{\kappa})$ is known and the value of $\boldsymbol{\Sigma}$ is fixed, it is possible to find the BLUP estimator for $\boldsymbol{\zeta}$ using the Newton-Raphson algorithm [Henderson (1975)], including the predictors for $\boldsymbol{U}$ and $\boldsymbol{V}$. Starting with $\boldsymbol{\zeta}_0$, an initial value for $\boldsymbol{\zeta}$, for $\ell = 1,2, \ldots$, the iterative procedure is based on

$$\boldsymbol{\zeta}_{\ell} = \boldsymbol{\zeta}_{\ell-1} + \boldsymbol{B}_{\ell}^{-1} \left(\frac{\partial l}{\partial \boldsymbol{\zeta}}\right)\Big|_{\boldsymbol{\zeta}=\boldsymbol{\zeta}_{\ell-1}}, \ \ell = 1,2, \ldots, \tag{7}$$

where $\boldsymbol{B}$is the negative of the the second derivative of l with respect to $\boldsymbol{\zeta}$. More details on this procedure are provided in Appendix A. Denote $\hat{\boldsymbol{\zeta}}$ as the estimator for $\boldsymbol{\zeta}$ given by (7) after convergence. Denote the partitioned matrix $\boldsymbol{B}$ and its inverse, evaluated at $\hat{\boldsymbol{\zeta}}$ according to the ordering $\boldsymbol{\gamma} \mid \boldsymbol{\beta} \mid \boldsymbol{U} \mid \boldsymbol{V}$, respectively, as

$$\widehat{\boldsymbol{B}}=\begin{pmatrix}\widehat{\boldsymbol{B}}_{11} & \widehat{\boldsymbol{B}}_{12} & \widehat{\boldsymbol{B}}_{13} & \widehat{\boldsymbol{B}}_{14}\\ \widehat{\boldsymbol{B}}_{21} & \widehat{\boldsymbol{B}}_{22} & \widehat{\boldsymbol{B}}_{23} & \widehat{\boldsymbol{B}}_{24}\\ \widehat{\boldsymbol{B}}_{31} & \widehat{\boldsymbol{B}}_{32} & \widehat{\boldsymbol{B}}_{33} & \widehat{\boldsymbol{B}}_{34}\\ \widehat{\boldsymbol{B}}_{41} & \widehat{\boldsymbol{B}}_{42} & \widehat{\boldsymbol{B}}_{43} & \widehat{\boldsymbol{B}}_{44}\end{pmatrix} \quad \text{and}$$

$$\widehat{\boldsymbol{B}}^{-1}=\begin{pmatrix}\widehat{\boldsymbol{A}}_{11} & \widehat{\boldsymbol{A}}_{12} & \widehat{\boldsymbol{A}}_{13} & \widehat{\boldsymbol{A}}_{14}\\ \widehat{\boldsymbol{A}}_{21} & \widehat{\boldsymbol{A}}_{22} & \widehat{\boldsymbol{A}}_{23} & \widehat{\boldsymbol{A}}_{24}\\ \widehat{\boldsymbol{A}}_{31} & \widehat{\boldsymbol{A}}_{32} & \widehat{\boldsymbol{A}}_{33} & \widehat{\boldsymbol{A}}_{34}\\ \widehat{\boldsymbol{A}}_{41} & \widehat{\boldsymbol{A}}_{42} & \widehat{\boldsymbol{A}}_{43} & \widehat{\boldsymbol{A}}_{44}\end{pmatrix}.$$

According to McGilchrist and Yau (1995) and Yau and Ng (2001), the REML estimate for $\boldsymbol{\Sigma}$ is then given by

$$\widehat{\boldsymbol{\Sigma}} = J^{-1}\mathrm{diag}\{\mathrm{tr}\widehat{\boldsymbol{A}}_{33} + \widehat{\boldsymbol{U}}^T\widehat{\boldsymbol{U}}, \mathrm{tr}\widehat{\boldsymbol{A}}_{44} + \widehat{\boldsymbol{V}}^T\widehat{\boldsymbol{V}}\}. \quad (8)$$

Moreover,

$$\mathrm{Var}\begin{pmatrix}\hat{\gamma}\\ \hat{\beta}\end{pmatrix} = \begin{pmatrix}\widehat{\boldsymbol{A}}_{11} & \widehat{\boldsymbol{A}}_{12}\\ \widehat{\boldsymbol{A}}_{21} & \widehat{\boldsymbol{A}}_{22}\end{pmatrix} \text{ and}$$

$$\mathrm{Var}(\widehat{\boldsymbol{\Sigma}}) = \begin{pmatrix}\widehat{\boldsymbol{\Sigma}}_{11} & \widehat{\boldsymbol{\Sigma}}_{12}\\ \widehat{\boldsymbol{\Sigma}}_{21} & \widehat{\boldsymbol{\Sigma}}_{22}\end{pmatrix},$$

with

$$\widehat{\boldsymbol{\Sigma}}_{11} = \sigma_u^{-4}\left(J - 2\sigma_u^{-2}\mathrm{tr}\,\widehat{\boldsymbol{A}}_{33}\right) + \sigma_u^{-8}\mathrm{tr}\left(\widehat{\boldsymbol{A}}_{33}^2\right),$$

$$\widehat{\boldsymbol{\Sigma}}_{22} = \sigma_v^{-4}\left(J - 2\sigma_v^{-2}\mathrm{tr}\,\widehat{\boldsymbol{A}}_{44}\right) + \sigma_v^{-8}\mathrm{tr}\left(\widehat{\boldsymbol{A}}_{44}^2\right) \text{ and}$$

$$\widehat{\boldsymbol{\Sigma}}_{12} = \sigma_u^{-4}\sigma_v^{-4}\mathrm{tr}\,(\widehat{\boldsymbol{A}}_{34}\widehat{\boldsymbol{A}}_{43}).$$

In order to deal with the unknown $S_0(t|\boldsymbol{\kappa})$, plug the REML estimator obtained for $\boldsymbol{\zeta}$ and $\boldsymbol{\Sigma}$ into l and denote the resulting profile likelihood for $\boldsymbol{\kappa}$ by l_p. Considering a bidimensional grid for $\boldsymbol{\kappa}$, values for λ and α maximizing l_p are obtained. The Newton-Raphson approach can also be used. Considering $\boldsymbol{\kappa}_0$ a initial value for $\boldsymbol{\kappa}$, the Newton-Raphson algorithm is based on

$$\boldsymbol{\kappa}_\ell = \boldsymbol{\kappa}_{\ell-1} + \boldsymbol{C}_\ell^{-1}\left(\frac{\partial l_p}{\partial \boldsymbol{\kappa}}\right)\Big|_{\boldsymbol{\kappa}=\boldsymbol{\kappa}_{\ell-1}}, \ \ell = 1,2,..., \quad (9)$$

where $\boldsymbol{C}$ is the negative of the second derivative of the log-likelihood l_p with respect to $\boldsymbol{\kappa}$. Details of such procedure are given in Appendix B. As pointed out in Yau and Ng (2001), the asymptotic variance of the limiting value $\widehat{\boldsymbol{\kappa}}$ in (9) is not given by the inverse of $\boldsymbol{C}$, because the expected values of the partial derivatives with respect to λ or α considering the other parameters as fixed are not zero. However, a resampling method as the jackknife [Miller (1974)] may be considered to compute the standard deviations for the estimator of $\boldsymbol{\kappa}$. After getting such estimates, it is necessary to restart the procedure to update the estimates for $\boldsymbol{\zeta}$ and $\boldsymbol{\Sigma}$, with new updated values, compute the updated estimate for $\boldsymbol{\kappa}$, and so on. In summary, to find estimates for $\boldsymbol{\zeta}$, $\boldsymbol{\Sigma}$ and $\boldsymbol{\kappa}$, one has to iterate the procedures described by equations (7), (8) and (9) until convergence.

Model diagnostics can be performed considering a graphical analysis based on the Cox-Snell residuals, defined as

$$r_{ij} = H\left(t_{ij}; \widehat{\boldsymbol{\psi}}, \boldsymbol{x}_{ij}, \boldsymbol{z}_{ij}\right), i = 1, \ldots, n_j; j = 1, \ldots, J. \quad (10)$$

where $H\left(\cdot; \widehat{\boldsymbol{\psi}}, \boldsymbol{x}_{ij}, \boldsymbol{z}_{ij}\right)$ is the cumulative risk function evaluated at the parameter estimates $\widehat{\boldsymbol{\psi}} = (\widehat{\boldsymbol{\zeta}}, \widehat{\boldsymbol{\Sigma}}, \widehat{\boldsymbol{\kappa}})$. If the model is correct, then r_{ij}, $j = 1, \ldots, J; i = 1, \ldots, n_j$, should behave as a censored sample from an exponential distribution with mean one. A disadvantage of such procedure is that if the proposed model is not adequate, the Cox-Snell residuals will not give any information on the source of inadequacy.

Application

Kalbfleisch and Prentice (2002) have presented a data set related to a trial to compare treatments for oropharynx carcinoma. A total of 195 subjects, with lesions in three different sites is considered: faucial arch, tonsillar fossa and pharyngeal tongue. Each patient was enrolled in one of six participating clinics and randomly assigned to one of the treatments: radiation therapy and radiation therapy combined with chemotherapy. Information on some covariates (gender, tumor stage classification (T-stage), lymph node stage classification (N-stage), among others) which can have influence on the survival of the patients are also available. The aim of the study is to identify risk factors for cancer and compare the treatments with respect to survival prognostic. Patients undergone the treatment for a period of 90 days; after that they received post-treatments. Since there were no restrictions, the clinics could use different approaches for the post-treatment, which could have some influence on patients survival times as well as on the cure rates. This also may lead to some correlation between observed quantities such as the survival times for patients belonging to same clinic.

In Kalbfleisch and Prentice (2002) the clinics were considered as fixed effects. Consider now the proposed approach following the same lines as in Yau and Ng (2001) and Lopes and Bolfarine (2012), so that only patients with cancer located in the pharyngeal were considered, corresponding to a total of 66 patients, out of which 47 died and 19 were censored (29%). Variable T-stage measuring the tumor size is grouped in 4 categories: T_1, T_2 and T_3 which refer to primary tumor measurement and T_4 to massive invasive tumor. For the i–th individual in the j–th clinic, define the x_{ij} as

$$x_{ij} = \begin{cases} 1, \text{if tumor is classified as } T_1, T_2 \text{ or } T_3, \\ 0, \text{if tumor is classified as } T_4. \end{cases}$$

Failure is defined as death due to the tongue or oropharynx carcinoma. Deaths related to other causes were considered as censored data. Previous analyses [Kalbfleisch and Prentice (2002), Yau and Ng (2001), Lai and Yau (2008) and Lopes and Bolfarine (2012)] agree that there is strong evidence of a cure fraction in this data set. The proposed model in Section 3 is considered with

- $\boldsymbol{X} = \left(x_{ij}\right)_{n^* \times 1}$, $j = 1, \ldots, 6; i = 1, \ldots, n_j$; $n^* = \sum_{j=1}^{J} n_j$,
- $\boldsymbol{Z} = (\mathbf{1}_{n^*}, \boldsymbol{X})$, where $\mathbf{1}_{n^*}$ is a n-column vector of ones,
- $\boldsymbol{\beta} = (\beta_0, \beta_1)^{\boldsymbol{T}}$,
- $\boldsymbol{\gamma} = \gamma$, and
- $n_1 = 9, n_2 = 14, n_3 = 16, n_4 = 7, n_5 = 11, n_6 = 9, n^* = 66$.

Note that the same set of covariates to model the cure fraction and survival of non-cured individuals has been considered. Additionally, the cure rate of Chen et al. (1999) (i.e., with no random effects) based on (1) and an ordinary Weibull survival model were also fitted, that is, neither one took into account the effect of clinic.

Results of model fitting for the models described above are presented in Table 1. In addition, random effects estimates for each one of the clinics are shown in Table 2.

TABLE 1 ML AND REML ESTIMATES FOR THE MODEL PARAMETERS (SE).

Parameter	PTCRM with random effects (REML)	Independent PTCRM (ML)
γ	0.528 (0.207)	0.422 (0.375)
β_0	0.275 (0.039)	0.240(0.244)
β_1	0.441 (0.156)	0.520(0.409)
σ_u^2	0.003 (0.028)	-
σ_v^2	0.029 (0.014)	-
λ	0.00006 (0.00009)	0.0025 (0.00056)
α	1.478 (0.281)	1.472 (0.374)

Parameter	Independent Single Weibull (ML)
γ	0.573(0.177)
β_0	-
β_1	-
σ_u^2	-
σ_v^2	-
λ	0.00016 (0.00014)
α	1.073 (0.134)

TABLE 2 BLUP PREDICTORS FOR THE RANDOM EFFECTS.

Clinic	U_j	V_j
1	0.0020	-0.0419
2	0.0010	0.1043
3	-0.0096	-0.0861
4	0.0025	-0.0294
5	0.0020	0.0498
6	0.0022	0.0033

The figures in the first row of Table 1 suggest that when the cure fraction is excluded from the analysis, tumor stage seems to have a significant effect on the survival of susceptible patients. On the other hand, when a cure fraction model without random effects is considered, tumor stage no longer seems to be significant. In contrast, the results presented in Yau and Ng (2001) and Lai and Yau (2008) considering the mixture model have suggested that this variable is significant. For the random effect model considered here (first column of the table), tumor stage seems also to be significant on the cure fraction. In addition, the values for random effects shown in Table 2 are close to zero and, probably, non-significant given the estimates of their variances. However, a formal test for the hypothesis $H_0: \sigma_u^2 = \sigma_v^2 = 0$ to be developed to allow a more reliable assessment will be considered in a future work.

Point estimates for the cure rates $e^{-\theta_{ij}}$ for patients with small primary tumor $(X = 0)$ and a massive tumor $(X = 1)$ are presented in Table 3. Inspecting the values under an exploratory perspective, it seems that Clinic 3 is the one performing the best among all the clinics, so that in average, the activation time of the cells for susceptible individuals is higher. The estimated hazard ratio is $e^{\hat{\gamma}} = 1.70$, so the cells of non-cured subjects with a massive tumor have a 70% higher risk to cause death when compared to non-cured patients with tumors in other stages.

TABLE 3 ESTIMATED CURE RATE FOR THE SIX CLINICS.

	Estimate cure rate	
Clinic	$Z = 0$	$Z = 1$
1	0.28	0.14
2	0.23	0.10
3	0.30	0.15
4	0.28	0.14
5	0.25	0.12
6	0.27	0.13

Cox-Snell's residuals were computed to check model fitting, considering (10). The cumulative risk function for those residuals was computed and compared with the cumulative distribution for the unitary exponential distribution. As it can be depicted from Figure 1, the two cumulative risk functions are close, suggesting a

reasonable model fit.

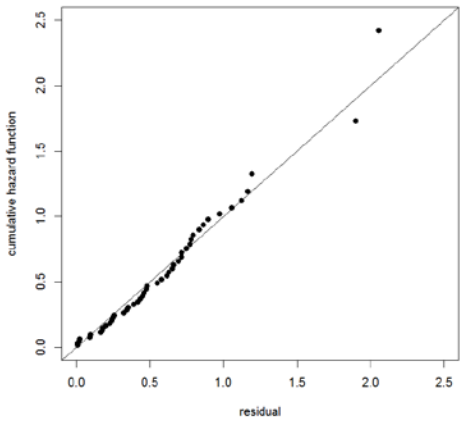

FIG. 1 CUMULATIVE HAZARD FUNCTIONS FOR COX-SNELL RESIDUALS AND EXPONENTIAL DISTRIBUTION (SOLID LINE).

Simulation

The results of a simulation study performed for the promotion time cure rate model with random effects have been reported so that results are comparable with the simulation study performed by Yau and Ng (2001). 10 clinics each with 15 patients are taken into consideration. Patients are randomly assigned to the treatment group ($x_{ij} = 1$) and the remaining patients are assigned to the control group ($x_{ij} = 0$), for each clinic. Following the simulation in Yau and Ng (2001), the same covariate for the cure fraction with the addition of an intercept was in consideration as well, i.e., $z_{ij} = [1, x_{ij}]$. Values of N_{ij} are simulated from the Poisson distribution with mean θ_{ij}. The values of θ_{ij} are given as in (6) and the V_j are generated from $N(0, \sigma_v^2)$. For cured individuals, that is, $N_{ij} = 0$, the failure time is infinity and for non-cured individuals, random variables $W_{1j}, \dots, W_{N_{ij}\,j}$ representing promotion times, with density given by $f_0(t, x_{ij}) = h_0(t|\boldsymbol{\kappa})\exp\{x_{ij}\gamma + U_j\}S_0(t|\boldsymbol{\kappa})^{\exp\{x_{ij}\gamma + U_j\}}$ have been generated, with U_j generated from the $N(0, \sigma_u^2)$ and the minimum values of these survival times considered to be failure times. If the survival time is greater than the censoring time C, then it is considered to be a censored observation. In the simulation study, $S_0(t|\boldsymbol{\kappa})$ is Weibull as in (4) is regarded, with parameters λ and α. We set $\lambda = 0.01$ was set in all the simulations and 3 groups for parameters γ, β_0, β_1 were in consideration. For each group simulations were performed with $\{\alpha = 1, \sigma_u^2 = 0.5, \sigma_v^2 = 0.5, C = 500\}$, and the effect of changes on $\alpha, \sigma_u^2, \sigma_v^2$ and C was studied. For each combination, 500 simulations were generated.

We expect that

- As α varies from 1 to 1.3, the performance of the estimators is improved since the W_{ij} will have smaller expectation making it more probable that $\min\{W_{ij},\ i = 1, \dots, N_{ij}\}$ is smaller than C allowing an easier identification on cured patients.
- As the variances of the random effects grow from 0.5 to 1 and the censoring times is reduced from 500 to 300, estimators performance will be worse since the variances of the random effects are increased (and consequently the variability of the generated failure amplifies because the censoring time will increase in the group of susceptible individuals.

TABLE 4 ESTIMATED BIASES FOR 500 SIMULATIONS OF REML ESTIMATION.

Parameter	True value	Average bias	SE_1	SE_2
(a) Sim. Set I ($\lambda = 0.01, \gamma = -0.5, \beta_0 = 0, \beta_1 = 0.5$)				
Sim. 1 ($\alpha = 1, \sigma_u^2 = 0.5, \sigma_v^2 = 0.5, C = 500$)				
γ	-0.5	0.003	0.367	0.403
β_0	0.0	-0.004	0.265	0.284
β_1	0.5	0.015	0.299	0.314
σ_u^2	0.5	-0.022	0.327	0.397
σ_v^2	0.5	0.034	0.319	0.331
Sim. 2 ($\alpha = 1.3, \sigma_u^2 = 0.5, \sigma_v^2 = 0.5, C = 500$)				
γ	-0.5	0.005	0.273	0.290
β_0	0.0	0.032	0.259	0.273
β_1	0.5	-0.001	0.244	0.240
σ_u^2	0.5	-0.006	0.306	0.322
σ_v^2	0.5	0.010	0.301	0.312
Sim. 3 ($\alpha = 1, \sigma_u^2 = 1, \sigma_v^2 = 1, C = 500$)				
γ	-0.5	0.003	0.406	0.443
β_0	0.0	0.004	0.352	0.344
β_1	0.5	0.011	0.331	0.335
σ_u^2	1.0	-0.005	0.600	0.766
σ_v^2	1.0	0.062	0.594	0.713
Sim. 4 ($\alpha = 1, \sigma_u^2 = 0.5, \sigma_v^2 = 0.5, C = 300$)				
γ	-0.5	-0.036	0.553	0.565
β_0	0.0	0.016	0.271	0.265
β_1	0.5	0.056	0.446	0.426
σ_u^2	0.5	0.030	0.391	0.458
σ_v^2	0.5	0.051	0.338	0.376
(b) Sim. Set II ($\lambda = 0.01, \gamma = -0.5, \beta_0 = -0.5, \beta_1 = 0.5$)				
Sim. 5 ($\alpha = 1, \sigma_u^2 = 0.5, \sigma_v^2 = 0.5, C = 500$)				
γ	-0.5	0.006	0.383	0.426
β_0	-0.5	0.019	0.272	0.277
β_1	0.5	-0.009	0.308	0.314
σ_u^2	0.5	0.011	0.360	0.447
σ_v^2	0.5	0.022	0.326	0.362
Sim. 6 ($\alpha = 1.3, \sigma_u^2 = 0.5, \sigma_v^2 = 0.5, C = 500$)				
γ	-0.5	0.010	0.275	0.291
β_0	-0.5	0.029	0.268	0.276
β_1	0.5	-0.014	0.253	0.250
σ_u^2	0.5	0.027	0.333	0.343
σ_v^2	0.5	0.010	0.313	0.328

TABLE 4 ESTIMATED BIASES FOR 500 SIMULATIONS OF REML ESTIMATION (CONTINUATION).

Parameter	True value	Average bias	SE_1	SE_2
(b) Sim. Set II ($\lambda = 0.01, \gamma = -0.5, \beta_0 = -0.5, \beta_1 = 0.5$)				
Sim. 7 ($\alpha = 1, \sigma_u^2 = 1, \sigma_v^2 = 1, C = 500$)				
γ	-0.5	0.012	0.399	0.438
β_0	-0.5	0.022	0.354	0.373
β_1	0.5	-0.020	0.318	0.322
σ_u^2	1.0	0.030	0.630	0.847
σ_v^2	1.0	0.021	0.586	0.686
Sim. 8 ($\alpha = 1, \sigma_u^2 = 0.5, \sigma_v^2 = 0.5, C = 300$)				
γ	-0.5	-0.059	0.618	0.618
β_0	-0.5	0.003	0.278	0.270
β_1	0.5	0.051	0.484	0.441
σ_u^2	0.5	0.112	0.456	0.516
σ_v^2	0.5	0.028	0.344	0.388
(c) Sim. Set III ($\lambda = 0.01, \gamma = -1, \beta_0 = -1, \beta_1 = 1$)				
Sim. 9 ($\alpha = 1, \sigma_u^2 = 0.5, \sigma_v^2 = 0.5, C = 500$)				
γ	-1.0	-0.095	0.753	0.704
β_0	-1.0	0.018	0.285	0.288
β_1	1.0	0.101	0.634	0.533
σ_u^2	0.5	0.080	0.429	0.530
σ_v^2	0.5	-0.003	0.336	0.373
Sim. 10 ($\alpha = 1.3, \sigma_u^2 = 0.5, \sigma_v^2 = 0.5, C = 500$)				
γ	-1.0	0.019	0.282	0.278
β_0	-1.0	0.062	0.285	0.288
β_1	1.0	-0.006	0.269	0.274
σ_u^2	0.5	-0.007	0.333	0.356
σ_v^2	0.5	0.019	0.329	0.357
Sim. 11 ($\alpha = 1, \sigma_u^2 = 1, \sigma_v^2 = 1, C = 500$)				
γ	-1.0	-0.033	0.566	0.614
β_0	-1.0	0.046	0.364	0.359
β_1	1.0	0.026	0.454	0.443
σ_u^2	1.0	0.034	0.669	0.928
σ_v^2	1.0	0.009	0.602	0.715
Sim. 12 ($\alpha = 1, \sigma_u^2 = 0.5, \sigma_v^2 = 0.5, C = 300$)				
γ	-1.0	-0.120	1.045	0.785
β_0	-1.0	0.020	0.289	0.310
β_1	1.0	0.163	0.870	0.622
σ_u^2	0.5	0.242	0.566	0.632
σ_v^2	0.5	-0.020	0.353	0.410

Simulation results are presented in Table 4. SE_1 and SE_2 represent, respectively, the mean of the standard deviations and the standard deviations of the estimates for the 500 simulated samples. The table reveals that the bias corresponding to the parameters related to the cure fraction, namely β_0 and β_1, is small except for the case $C = 300$ for which the bias of β_1 is substantially increased, specially for parameter groups II and III. Concerning the standard deviation, it can be noticed that SE_1 is slightly smaller than SE_2 indicating that the standard deviations of β_0 and β_1 are slightly underestimated. One exception occurs in the parameter group III ($C = 300$) for which the standard deviation was considerably underestimated. Regarding γ, it can be noticed that the bias is small except for group III and $C = 300$. The standard deviation of γ is slightly underestimated, except for parameter group III ($C = 300$) for which the standard deviation is substantially overestimated.

For the random effects it can be depicted that the bias is small, an exception being the case when σ_u^2 and σ_v^2 increase from 0.5 to 1.0 and C decreases from 500 to 300, in which case the bias increases, although the bias of σ_v^2 is always smaller than that of σ_u^2. The variances are slightly underestimated in all parameter groups, except in the case where σ_u^2 and σ_v^2 increase and the censoring variable diminishes. It is also worth noticing that as α increases from 1.0 to 1.3, the performance (evaluated in terms of standard deviation) of all estimators improves.

Final Discussion

This paper extends the cure rate model in Lopes and Bolfarine (2012) by considering random effects in both, the cure probability and the survival function for individuals that are at risk. The random effects are considered to be normally distributed. An approach to model adequacy is developed defining Cox-Snell type residuals for the proposed model. The model is parametrized in terms of the cured fraction which is then linked to covariates. The approach used for parameter estimation is the restricted maximum likelihood (REML) approach proposed in McGilchrist and Yau (1995). Simulation studies are performed and results of a real data analysis are presented indicating good performance of the proposed approach. Moreover, residual analysis performed for the model considered to analyze the real data has indicated that the methodology considered in this paper is suitable to deal with real data sets of the considered type. An important case exclusive from consideration in this work is the situation where the random effects U_j and V_j are correlated, which could be the case in many practical applications. This generalization is the subject of a forthcoming paper.

Appendix A: The BLUP Estimation Procedure

The score and matrix $\boldsymbol{B}$, computed based on the second derivative of l with respect to $\boldsymbol{\zeta}$, are as follows:

$$\frac{\partial l}{\partial \boldsymbol{\zeta}} = \begin{bmatrix} \boldsymbol{X}' & 0 \\ 0 & \boldsymbol{Z}' \\ \boldsymbol{G}' & 0 \\ 0 & \boldsymbol{G}' \end{bmatrix} \begin{bmatrix} \partial l_1/\partial \boldsymbol{\xi} \\ \partial l_1/\partial \boldsymbol{\phi} \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ \sigma_u^{-2}\boldsymbol{U} \\ \sigma_v^{-2}\boldsymbol{V} \end{bmatrix},$$

$$\boldsymbol{B} = \begin{bmatrix} \boldsymbol{X}' & 0 \\ 0 & \boldsymbol{Z}' \\ \boldsymbol{G}' & 0 \\ 0 & \boldsymbol{G}' \end{bmatrix} \begin{bmatrix} -\partial^2 l_1/\partial\boldsymbol{\xi}\partial\boldsymbol{\xi}' & -\partial^2 l_1/\partial\boldsymbol{\xi}\partial\boldsymbol{\phi}' \\ -\partial^2 l_1/\partial\boldsymbol{\phi}\partial\boldsymbol{\xi}' & -\partial^2 l_1/\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}' \end{bmatrix} \begin{bmatrix} \boldsymbol{X} & 0 & \boldsymbol{G} & 0 \\ 0 & \boldsymbol{Z} & 0 & \boldsymbol{G} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_u^{-2} I_J & 0 \\ 0 & 0 & 0 & \sigma_v^{-2} I_J \end{bmatrix},$$

with $\boldsymbol{\xi} = \boldsymbol{X}'\boldsymbol{\gamma} + \boldsymbol{G}'\boldsymbol{U}$ and $\boldsymbol{\phi} = \boldsymbol{Z}'\boldsymbol{\beta} + \boldsymbol{G}'\boldsymbol{V}$. I_J is the identity matrix of dimension J and $\boldsymbol{X}, \boldsymbol{Z}$ and $\boldsymbol{G}$ are the design matrices of $\boldsymbol{\gamma}, \boldsymbol{\beta}$ and the random effects $\boldsymbol{U}$ and $\boldsymbol{V}$ respectively. Let $H_0(t|\boldsymbol{\kappa})$ be the baseline cumulative hazard. The first and second partial derivatives of l_1 with respect to the elements of $\boldsymbol{\xi}$ and $\boldsymbol{\phi}$ are given as follows:

$$\frac{\partial l_1}{\partial \xi_{ij}} = \delta_{ij}\left(1 - H_0(t_{ij}|\boldsymbol{\kappa})e^{\xi_{ij}}\right), \quad -e^{\phi_{ij}+\xi_{ij}} H_0(t_{ij}|\boldsymbol{\kappa})\left[S_0(t_{ij}|\boldsymbol{\kappa})\right]^{\exp\,\xi_{ij}},$$

$$-\frac{\partial^2 l_1}{\partial \xi_{ij}^2} = H_0(t_{ij}|\boldsymbol{\kappa})e^{\xi_{ij}}\left[\delta_{ij} + e^{\phi_{ij}}\left[S_0(t_{ij}|\boldsymbol{\kappa})\right]^{\exp\,\xi_{ij}}\left(1 - H_0(t_{ij}|\boldsymbol{\kappa})e^{\xi_{ij}}\right)\right],$$

$$\frac{\partial l_1}{\partial \phi_{ij}} = \delta_{ij} - e^{\phi_{ij}}\left(1 - \left[S_0(t_{ij}|\boldsymbol{\kappa})\right]^{\exp\,\xi_{ij}}\right),$$

$$-\frac{\partial^2 l_1}{\partial \phi_{ij}^2} = e^{\phi_{ij}}\left(1 - \left[S_0(t_{ij}|\boldsymbol{\kappa})\right]^{\exp\,\xi_{ij}}\right),$$

$$-\frac{\partial^2 l_1}{\partial \phi_{ij}\,\partial \xi_{ij}} = e^{\phi_{ij}+\xi_{ij}} H_0(t_{ij}|\boldsymbol{\kappa})\left[S_0(t_{ij}|\boldsymbol{\kappa})\right]^{\exp\,\xi_{ij}},$$

$$-\frac{\partial^2 l_1}{\partial \xi_{ij}\,\partial \xi_{i'j'}} = 0, \qquad -\frac{\partial^2 l_1}{\partial \phi_{ij}\,\partial \xi_{i'j'}} = 0,$$

$$-\frac{\partial^2 l_1}{\partial \phi_{ij}\,\partial \xi_{i'j'}} = 0, \qquad \text{for } i \neq i' \text{ or } j \neq j'.$$

Appendix B: Estimation of Parameters of Baseline Hazard

For a Weibull baseline hazard, the cumulative hazard is $H_0(t|\boldsymbol{\kappa}) = \lambda t^{\alpha}$. Let $S_{ij} = \left[S_0(t_{ij}|\boldsymbol{\kappa})\right]^{\exp\,\xi_{ij}}$ and $\theta_{ij} = e^{\phi_{ij}}$. The first and second partial derivatives of l_1 with respect to the parameters of the baseline hazard and matrix are given as follows

$$\frac{\partial l_p}{\partial \lambda} = \sum_{j=1}^{J}\sum_{i=1}^{n_j}\left\{\delta_{ij}\left(\lambda^{-1} - e^{\xi_{ij}} t_{ij}^{\alpha}\right) - \theta_{ij}\, e^{\xi_{ij}} t_{ij}^{\alpha} S_{ij}\right\},$$

$$-\frac{\partial^2 l_p}{\partial \lambda^2} = \sum_{j=1}^{J}\sum_{i=1}^{n_j}\left\{\delta_{ij}\,\lambda^{-2} - \theta_{ij}\left(e^{\xi_{ij}} t_{ij}^{\alpha}\right)^2 S_{ij}\right\},$$

$$\frac{\partial l_p}{\partial \alpha} = \sum_{j=1}^{J}\sum_{i=1}^{n_j}\left\{\delta_{ij}\left(\alpha^{-1} + \log t_{ij} - \lambda e^{\xi_{ij}} t_{ij}^{\alpha}\log t_{ij}\right) - \lambda\theta_{ij}\, e^{\xi_{ij}} t_{ij}^{\alpha} S_{ij}\log t_{ij}\right\},$$

$$-\frac{\partial^2 l_p}{\partial \alpha^2} = \sum_{j=1}^{J}\sum_{i=1}^{n_j}\left\{\delta_{ij}\left(\alpha^{-2} + \lambda e^{\xi_{ij}} t_{ij}^{\alpha}\log^2 t_{ij}\right) + \lambda\theta_{ij}\, e^{\xi_{ij}} t_{ij}^{\alpha} S_{ij}\log^2 t_{ij}\left(1 - \lambda e^{\xi_{ij}} t_{ij}^{\alpha}\right)\right\},$$

$$-\frac{\partial^2 l_p}{\partial \alpha\,\partial \lambda} = \sum_{j=1}^{J}\sum_{i=1}^{n_j} e^{\xi_{ij}} t_{ij}^{\alpha}\log t_{ij}\left\{\delta_{ij} + \theta_{ij} S_{ij}\left(1 - \lambda e^{\xi_{ij}} t_{ij}^{\alpha}\right)\right\},$$

ACKNOWLEDGMENT

The first author acknowledges the financial support from CONICYT-Chile. The authors also want to thank the referees whose valuable comments improved the presentation of this paper.

REFERENCES

Berkson, J., and Gage, R. "Survival curve for cancer patients following treatment." Journal of the American Statistical Association 47 (1952): 501-515.

Chen, M.H., and Ibrahim, J.G., Sinha, D. "A new Bayesian model for survival data with a surviving fraction." Journal of the American Statistical Association 94 (1999): 909-919.

Davis, P., and Rabinowitz, P. Methods of numerical integration. New York: Academic Press, 1975.

Goldman, A. "Survivorship analysis when cure is a possibility: a Monte Carlo study." Statistics in Medicine 3 (1984): 153-163.

Henderson, C.R. "Best linear unbiased estimation and prediction under a selection model." Biometrics 31 (1975): 423-447.

Kalbfleisch, J.D., and Prentice, R.L. The Statistical Analysis of Failure Data. New York: Wiley, 2002.

Kaplan, E.L., Meier, P. "Nonparametric estimation from incomplete observations." Journal of the American Statistical Association 53 (1958): 457-481.

Lai, X., and Yau, K. "Long-term survivor model with bivariate random effects: Applications to bone marrow transplant and carcinoma study data." Statistics in Medicine 27 (2008): 5692-5708.

Lopes, C.M., and Bolfarine, H. (2012). "Random effects in promotion time cure rate models." Computational

Statistics and Data Analysis 56 (2012): 75-87.

Maller, R.A., and Zhou, S. (1996). Survival analysis with long term survivors. New York: Wiley, 1996.

McGilchrist, C.A., and Yau, K.K. "The derivation of BLUP, ML, REML estimation methods for generalised linear mixed models." Communications in Statistics - Theory and Methods 24 (1995): 2963-2980.

Miller, R.G. "The jackknife: a review." Biometrika 61 (1974): 1-15.

Mizoi, M.F., Bolfarine, H., and Lima, A.C.P. (2007). "Cure rate models with measurement errors." Communications in Statistics - Simulation and Computation 36(1) (2007): 185-196.

Rodrigues, J., Cancho, V.G., Castro, M.A., and Louzada-Neto, F. "On the unification of the long-term survival models." Statistics and Probability Letters 79 (2008): 753-759.

Sposto, R., Sather, H., and Baker, S. "A comparison of tests of the difference in the proportion of patients who are cured." Biometrics 48 (1992): 87-99.

Yakovlev, A.Y., and Tsodikov, A.D. Stochastic Models of Tumor Latency and their Biostatistical Applications. New Jersey: World Scientific, 1996.

Yau K., and Ng, A. "Long-term survivor mixture model with random effects: application to a multi-centre clinical trial of carcinoma." Statistics in Medicine 20 (2001):1591-1607.

Yin, G., and Ibrahim, J.G. (2005). Cure rate models: a unified approach. Canadian Journal of Statistics 33(4) (2005): 559-570.